



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/785,238	02/24/2004	Takeshi Otani	FUJR 20.949	1025
26304	7590	09/26/2008	EXAMINER	
KATTEN MUCHIN ROSENMAN LLP 575 MADISON AVENUE NEW YORK, NY 10022-2585				COLUCCI, MICHAEL C
ART UNIT		PAPER NUMBER		
2626				
		MAIL DATE		DELIVERY MODE
		09/26/2008		PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/785,238	OTANI ET AL.	
	Examiner	Art Unit	
	MICHAEL C. COLUCCI	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on ____.
- 2a) This action is **FINAL**. 2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 31 and 32 is/are pending in the application.
 - 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) Claim(s) ____ is/are allowed.
- 6) Claim(s) 31 and 32 is/are rejected.
- 7) Claim(s) ____ is/are objected to.
- 8) Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on 24 February 2004 is/are: a) accepted or b) objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 - a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. ____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) <input type="checkbox"/> Notice of References Cited (PTO-892)	4) <input type="checkbox"/> Interview Summary (PTO-413)
2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)	Paper No(s)/Mail Date. ____ .
3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)	5) <input type="checkbox"/> Notice of Informal Patent Application
Paper No(s)/Mail Date ____ .	6) <input type="checkbox"/> Other: ____ .

DETAILED ACTION

NOTE: Examiner acknowledges the cancellation of claims 1-30.

Response to Arguments

1. Applicant's arguments filed 06/30/2008 have been fully considered but they are not persuasive.

Arguments 1 and 2 (page 5-7):

- “It is respectfully submitted that neither Borth nor Ananthaiyer teach a voice activity detection of claim 31 as shown above, whether used alone or in combination”
- “It is respectfully submitted that neither Borth nor Ananthaiyer teach the voice activity detection features of claim 32 as shown above, whether used alone or in combination. It is further submitted that the shortcomings of Borth and Ananthaiyer are not addressed by the relied upon portions of Sugar”

Response to arguments 1 and 2:

Examiner construes the flatness of a voice/noise signal to be functionally equivalent and equally effective to that of the power or energy level of a signal, wherein low energy corresponds to a flat signal and higher energy corresponds to a less flat signal, wherein voice will be present or not relative to high or low elevation/flatness in a voice signal. (Present invention Fig. 15)

Additionally, Examiner construes the finding of a maximum value of the frequency spectrum to be functionally equivalent and equally effective to revealing all the peaks in a signal and finding the maximum peak relative to the spectrum of a signal. (Present invention Fig. 11A and 11B).

Further, Examiner construes decoding to be necessary when conducting analog and digital conversion schemes.

Examiner takes the position that Borth in fact teaches the limitations of the present invention, wherein Borth teaches that an apparatus and method is provided for automatically performing background noise estimation for use with an acoustic noise suppression system, wherein the background noise from a noisy pre-processed input signal--the speech-plus-noise signal available at the input of the noise suppression system--is attenuated to produce a noise-suppressed post-processed output signal--speech-minus-noise signal provided at the output of the noise suppression system--by spectral gain modification. The automatic background noise estimator includes a noise estimation means which generates and stores an estimate of the background noise power spectral density based upon the pre-processed input signal. The background noise estimator of the present invention further includes a noise detection means, such as an energy valley detector, which performs the speech/noise decision based upon the post-processed signal energy level. The noise detection means

provides this speech/noise decision to the noise estimation means such that the background noise estimate is updated only when the detected minima of the post-processed signal energy is below a predetermined threshold. The novel technique of implementing post-processed speech energy for the noise detection means, thereby controlling the pre-processed speech energy to the noise estimation means, allows the present invention to generate a highly accurate background noise estimate for an acoustic noise suppression system (Col. 2 line 46 – Col. 3 line 6).

Further, Borth teaches (basic noise suppression system 100 implementing spectral gain modification as is well known in the art. A continuous time signal containing speech-plus-noise is applied to input 102 of the noise suppressor where it is then converted to digital form by analog-to-digital converter 105. This digital data is then segmented into blocks of data by the windowing operation (e.g., Hamming, Hanning, or Kaiser windowing techniques) performed by window 110. The choice of the window is similar to the choice of the filter response in an analog spectrum analysis. The noisy speech signal is converted into the frequency domain by Fast Fourier Transform (FFT) 115. The power spectrum of the noisy speech signal is then calculated by magnitude squaring operation 120, and applied to background noise estimator 125 and to power spectrum modifier 130 (Col. 3 lines 35-52).

Further, Borth also teaches a well known method using microphones, wherein Borth overcomes previous well known burdens of using multiple microphones limitation through the use of a single input of speech and noise together, wherein Borth teaches an estimate of the background noise is to implement a second microphone, located at a distance away from the user's first microphone, such that it picks up only background noise. This technique has been shown to provide a significant improvement in signal-to-noise ratio (SNR). However, it is very difficult to achieve the required isolation of the second microphone from the speech source while at the same time attempting to pick up the same background noise environment as them first microphone (Col. 1 lines 25-39).

Further, Borth teaches energy valley detector 440 utilizes the overall energy estimate from combiner 460 to detect the pauses in speech. This is accomplished in three steps. First, an initial valley level is established. If the background noise estimator has not previously been initialized, then an initial valley level is created by loading initialization value 455. Otherwise, the previous valley level is maintained as its post-processed background noise energy history. Next, the previous (or initialized) valley level is updated to reflect current background noise conditions. This is accomplished by comparing the previous valley level to the value of the single overall energy estimate from combiner 460. A current valley level is created by this updating process, which will be described in detail in FIG. 6b. The third step performed by energy valley detector 440 is

that of making the actual speech/noise decision. A preselected valley level offset, represented in FIG. 4 by valley offset 445, is added to the updated current valley level to produce a noise threshold level. Then the value of the single overall (post-processed) energy estimate is again compared, only this time to the noise threshold level. When this energy estimate is less than the noise threshold level, energy valley detector 440 generates a speech/noise control signal (valley detect signal) indicating that no voice is present. (Col. 7 lines 3-29).

Though Borth teaches summation from band pass filter outputs and spectral subtraction, Borth does not specifically teach finding a peak/max value and adding up differences between spectral components and the maximum value thereof, and generates resulting sum of the differences as the speech flatness factor wherein the flatness evaluator calculates an average of spectral components of the voice/noise data, normalizes the resulting sum of the differences by dividing by -the calculated average, and outputs a normalized voice/noise flatness factor. Therefore, the reference of Sugar has been introduced to further strengthen the prior art of Borth.

Sugar teaches a max peak detector (Fig. 6 item 210 "MaxPeak"), wherein a peak detector 210, as shown in FIG. 6, comprises a comparator 212, a register file 214, a FIFO 216 and a FIFO 218. The comparator 212 compares the dB power value PDB(k) with the peak threshold (SD_PEAKTH). The FIFO 216 stores a

data word that indicates which frequency bins k had a power value above the peak threshold, and which did not. For example, if the FFT outputs 256 FFT values, the FIFO 216 stores a 256 bit word, with 1's indicating FFT values that exceed the peak threshold and 0's indicating FFT values that do not exceed the peak threshold. The register file 214 stores the maximum peak power value in any set of contiguous FFT values that exceed the peak threshold. This maxpeak information is used in the pulse detector (Sugar [0069]).

Further, Sugar teaches computing Fast Fourier Transform (FFT) values at a plurality of frequency bins from a digital signal representing activity in a frequency band during a time interval; computing the power at each frequency bin; adding the power at each frequency bin for a current time interval with the power at the corresponding frequency bin for a previous time interval to obtain a running sum of the power at each frequency bin; comparing the power at each frequency bin with a power threshold to obtain a duty count of the number of times that the power at each frequency bin exceeds the power threshold over time intervals; and comparing the power at each frequency bin for a current time interval with the power at the corresponding frequency bin for a previous time interval to track the maximum power in each frequency bin over time intervals. This process may also be implemented by instructions encoded on a processor readable medium that, when executed by a processor, cause the processor to perform these same steps (Sugar [0231]).

Though the combined teaching of Borth in view of Sugar teaches max peak detection, summation and comparison of adjacent frames of spectral information, the combined teaching fails to teach calculating an average of spectral components of the voice/noise data, normalizes the resulting sum of the differences by dividing by the calculated average, and outputs a normalized voice/noise flatness factor. Therefore, the reference of Ananthaiyer has been introduced to further strengthen the prior art of Borth in view of Sugar.

Ananthaiyer teaches a voice detector that maintains an average difference of the minimum AMDF values AvgDiffAMDF which is a running sum of the differences between the minimum local AMDF value for the interval m and the minimum local AMDF value for the previous interval (m-1) (Ananthaiyer Col. 4 lines 43-48).

Further, Ananthaiyer teaches normalization through an update interval logic and decision interval logic of FIG. 7. A signal detector apparatus for characterizing a signal over a detection cycle i, the detection cycle i having a number of intervals, each interval having a predetermined number of input samples 650, the device comprising: first logic 654 for determining an Average Magnitude Difference Function (AMDF) value 652 for each of a predetermined range of pitch frequencies K over the intervals; second logic 656 for determining an average difference AMDF value over the intervals equal to the sum of the difference

between a first minimum AMDF value from each interval m and a second minimum AMDF value from each interval (m-1); third logic 658 for determining a minimum AMDF value over the intervals; fourth logic 660 for determining a sum of the AMDF values over the intervals; fifth logic 662 or computing a first metric equal to the minimum AMDF value over the intervals divided by the sum of the AMDF values over the intervals; sixth logic 664 for computing a second metric equal to the average difference AMDF value over the intervals divided by the sum of the AMDF values over the intervals; and seventh logic 666 for utilizing said first metric and said second metric to determine whether the signal is one of a noise signal, a tone signal, and a voice signal (Ananthaiyer Col. 8 liens 33-55).

Furthermore, Ananthaiyer teaches a noise and silence discerning operation and adjustable threshold value method of determining if the signal is a noise signal in step 404. In step 404, the signal is characterized as noise, and the logic proceeds to step 410, if any of a number of conditions is true. First, the signal is characterized as noise if the AMDF.sub.sum is equal to zero. This case represents the detection of absolute silence. Second, the signal is characterized as noise if the AMDF.sub.norm for the current detection cycle i is greater than a threshold N, representing a large value of AMDF.sub.norm. Finally, the signal is characterized as noise if the signal detected in the previous detection cycle (i-1) was noise and the AMDF.sub.norm is greater than a threshold N2N which is less stringent than N. This condition applies the rule from the first observed

characteristic described above, specifically that the threshold for detecting subsequent noise signals can be made less stringent (Ananthaiyer Col. 6 lines 39-56).

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

3. Claims 31 and 32 are rejected under 35 U.S.C. 103(a) as being unpatentable over Borth et al, US 4630304 (hereinafter Borth) in view of Sugar et al USPGPUB 20030198304 A1 (hereinafter Sugar) and further in view of Ananthaiyer et al US 6385548 B (hereinafter Ananthaiyer).

Re claim 31, Borth teaches voice activity detector that detects talkspurts in an input coded speech signal and an input voice/noise signal (Col. 2 line 46 - Col. 3 line 6), comprising:

a first input controller that comprises a signal receiver and a decoder, wherein the signal receiver supplies the input coded speech signal to the decoder and the decoder decodes the input coded speech signal into a decoded speech data (Col. 3 lines 35-52);

a second input controller that comprises a microphone and an A/D converter, wherein the microphone supplies the input voice/noise signal to the A/D converter and the A/D converter converts the input voice/noise signal into a voice/noise data of digital form (Col. 3 lines 35-52);

a frequency spectrum calculator that calculates speech frequency spectrum of the speech data and calculates voice/noise frequency spectrum of the voice/noise data (Col. 3 lines 35-52);

a flatness evaluator that calculates a speech flatness factor indicating flatness of the speech frequency spectrum and calculates a voice/noise flatness factor indicating flatness of the voice/noise frequency spectrum (Col. 7 lines 3-29); and

(al) determining whether the speech data contains a talkspurt, by comparing the speech flatness factor of the speech frequency spectrum with a first predetermined threshold (Col. 7 lines 3-29).,

(bl) determining whether the voice/noise data contains a talkspurt, by comparing the normalized voice/noise flatness factor of the voice/noise frequency spectrum with a second predetermined threshold (Col. 7 lines 3-29).

(a) wherein, when the speech frequency spectrum is chosen for calculating the speech flatness factor,

adding up differences between spectral components and the maximum value thereof, and generates resulting sum of the differences as the speech flatness factor.

generating a resulting sum of the differences as the voice/noise flatness factor, and wherein the flatness evaluator calculates an average of spectral components of the

voice/noise data, normalizes the resulting sum of the differences by dividing by the calculated average, and outputs a normalized voice/noise flatness factor

(b) wherein, when the voice/noise frequency spectrum is chosen for calculating the voice/noise flatness factor,

the flatness evaluator finds a maximum value of the voice/noise frequency spectrum, adds up differences between spectral components and the maximum value thereof, and generates resulting sum of the differences as the voice/noise flatness factor, and wherein the flatness evaluator calculates an average of spectral components of the voice/noise data, normalizes the resulting sum of the differences by dividing by - the calculated average, and outputs a normalized voice/noise flatness factor; a voice/noise discriminator, performing:

However, Borth fails to teach the flatness evaluator finds a maximum value of the speech/voice/noise frequency spectrum,

Sugar teaches a max peak detector (Fig. 6 item 210 "MaxPeak"), wherein a peak detector 210, as shown in FIG. 6, comprises a comparator 212, a register file 214, a FIFO 216 and a FIFO 218. The comparator 212 compares the dB power value PDB(k) with the peak threshold (SD_PEAKTH). The FIFO 216 stores a data word that indicates which frequency bins k had a power value above the peak threshold, and which did not. For example, if the FFT outputs 256 FFT values, the FIFO 216 stores a 256 bit word, with 1's indicating FFT values that exceed the peak threshold and 0's indicating FFT values that do not exceed the peak threshold. The register file 214 stores the maximum

peak power value in any set of contiguous FFT values that exceed the peak threshold. This maxpeak information is used in the pulse detector (Sugar [0069]).

Further, Sugar teaches computing Fast Fourier Transform (FFT) values at a plurality of frequency bins from a digital signal representing activity in a frequency band during a time interval; computing the power at each frequency bin; adding the power at each frequency bin for a current time interval with the power at the corresponding frequency bin for a previous time interval to obtain a running sum of the power at each frequency bin; comparing the power at each frequency bin with a power threshold to obtain a duty count of the number of times that the power at each frequency bin exceeds the power threshold over time intervals; and comparing the power at each frequency bin for a current time interval with the power at the corresponding frequency bin for a previous time interval to track the maximum power in each frequency bin over time intervals. This process may also be implemented by instructions encoded on a processor readable medium that, when executed by a processor, cause the processor to perform these same steps (Sugar [0231]).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Borth to incorporate a flatness evaluator that finds a maximum value of the speech/voice/noise frequency spectrum as taught by Sugar because to allow for the determination of all peaks present in a signal in the frequency spectrum relevant to side by side frame based spectral values, wherein a running sum of the power at each frequency bin is obtained, comparing the power at each frequency bin with a power threshold to obtain a duty count of the number of times

that the power at each frequency bin exceeds the power threshold over time intervals relevant to previous framed spectral data (i.e. adjacent) (Sugar [0231]).

However, Borth in view of Sugar fails to teach adding up differences between spectral components and the maximum value thereof, and generates resulting sum of the differences as the speech flatness factor.

Generating a resulting sum of the differences as the voice/noise flatness factor, and wherein the flatness evaluator calculates an average of spectral components of the voice/noise data, normalizes the resulting sum of the differences by dividing by the calculated average, and outputs a normalized voice/noise flatness factor.

Ananthaiyer teaches a voice detector that maintains an average difference of the minimum AMDF values AvgDiffAMDF which is a running sum of the differences between the minimum local AMDF value for the interval m and the minimum local AMDF value for the previous interval (m-1) (Ananthaiyer Col. 4 lines 43-48).

Further, Ananthaiyer teaches normalization through an update interval logic and decision interval logic of FIG. 7. A signal detector apparatus for characterizing a signal over a detection cycle i, the detection cycle i having a number of intervals, each interval having a predetermined number of input samples 650, the device comprising: first logic 654 for determining an Average Magnitude Difference Function (AMDF) value 652 for each of a predetermined range of pitch frequencies K over the intervals; second logic 656 for determining an average difference AMDF value over the intervals equal to the sum of the difference between a first minimum AMDF value from each interval m and a

second minimum AMDF value from each interval (m-1); third logic 658 for determining a minimum AMDF value over the intervals; fourth logic 660 for determining a sum of the AMDF values over the intervals; fifth logic 662 or computing a first metric equal to the minimum AMDF value over the intervals divided by the sum of the AMDF values over the intervals; sixth logic 664 for computing a second metric equal to the average difference AMDF value over the intervals divided by the sum of the AMDF values over the intervals; and seventh logic 666 for utilizing said first metric and said second metric to determine whether the signal is one of a noise signal, a tone signal, and a voice signal (Ananthaiyer Col. 8 liens 33-55).

Furthermore, Ananthaiyer teaches a noise and silence discerning operation and adjustable threshold value method of determining if the signal is a noise signal in step 404. In step 404, the signal is characterized as noise, and the logic proceeds to step 410, if any of a number of conditions is true. First, the signal is characterized as noise if the AMDF.sub.sum is equal to zero. This case represents the detection of absolute silence. Second, the signal is characterized as noise if the AMDF.sub.norm for the current detection cycle i is greater than a threshold N, representing a large value of AMDF.sub.norm. Finally, the signal is characterized as noise if the signal detected in the previous detection cycle (i-1) was noise and the AMDF.sub.norm is greater than a threshold N2N which is less stringent than N. This condition applies the rule from the first observed characteristic described above, specifically that the threshold for detecting subsequent noise signals can be made less stringent (Ananthaiyer Col. 6 lines 39-56).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Borth in view of Sugar to incorporate adding up differences between spectral components and the maximum value thereof, and generates resulting sum of the differences as the speech flatness factor and generating a resulting sum of the differences as the voice/noise flatness factor, and wherein the flatness evaluator calculates an average of spectral components of the voice/noise data, normalizes the resulting sum of the differences by dividing by the calculated average, and outputs a normalized voice/noise flatness factor as taught by Ananthaiyer to allow for an average difference AMDF value over a group of intervals divided by the sum of the AMDF values over the intervals; and a first metric and second metric to determine whether the signal is one of a noise signal, a tone signal, and a voice signal, wherein noise if further classified and double checked to be purely noise or tonal where absolute silence and speech may be ruled out through a redundant method of threshold comparison (Ananthaiyer Col. 8 liens 33-55).

Claim 32 has been rejected with respect to claim 31, wherein claim 32 contains all its limitations within claim 31, wherein claim 31 distinguishes from claim 32 with the additional limitation of calculating a *maximum value* prior to summing differences, which was already addressed in the rejection of claim 31.

Conclusion

4. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael C. Colucci whose telephone number is (571)-270-1847. The examiner can normally be reached on 9:30 am - 6:00 pm, Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571)-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Michael C Colucci/
Examiner, Art Unit 2626
Patent Examiner
AU 2626
(571)-270-1847
Michael.Colucci@uspto.gov

/Richemond Dorvil/
Supervisory Patent Examiner, Art Unit 2626